THIS PAGE BLANK (USPTO)

OM protein - nucleic search, using frame_plus_p2n model

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

Run on: November 9, 2002, 04:40:52 ; Search time 2202 Seconds
                                    1397.431 Million cell updates/sec

Title: US-09-895-298A-83
Perfect score: 1002
Sequence: 1 MMNFQPPSKAWRASQMMTFF.........HDGSLDLRSRSVQEGNPRA 190

Scoring table: BLOSUM62
                Xgapop 10.0 , Xgapext 0.5
                Ygapop 10.0 , Ygapext 0.5
                Fgapop 6.0 , Fgapext 7.0
                Delop   6.0 , Delext   7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09895298/runat_06112002_160415_2344/app_query.fasta_1.327
-DB-EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09895298_@CGN_1_1.1716_@runat_06112002_160415_2344 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :   EST:*
                 1: em_estba:*
                 2: em_esthum:*
                 3: em_estin:*
                 4: em_estmu:*
                 5: em_estov:*
                 6: em_estpl:*
                 7: em_estro:*
                 8: em_htc:*
                 9: em_est1:*
                10: gb_est2:*
                11: gb_htc:*
                12: gb_est3:*
                13: gb_est4:*
                14: gb_est5:*
                15: em_estfun:*
                16: em_estom:*
                17: em_estom:*
                18: gb_gss:*
                19: em_gss_inv:*
                20: em_gss_pln:*
                21: em_gss_vrt:*
                22: em_gss_fun:*
                23: em_gss_mam:*
                24: em_gss_mus:*
                25: em_gss_other:*
                26: em_gss_pro:*
                27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1002	100.0	807	12	BG184448	BG184448
2	991	98.9	777	12	BG217799	BG217799
3	970.5	96.9	691	10	AW582253	AW582253 QV4-ST021
4	957	95.5	768	12	BG194960	BG194960 RST14018
5	923	92.1	789	12	BG221275	BG221275 RST41079
6	914	91.2	837	12	BG216071	BG216071 RST35885
7	906	90.4	823	12	BG183902	BG183902 RST2815 A
8	902	90.0	810	12	BG217267	BG217267 RST36969
9 c	857	85.5	2721	11	AK016573	AK016573 Mus muscu
10	799.5	79.8	6928	12	BF375244	BF375244 QV4-ST021
11	775	77.3	822	12	BG219307	BG219307 RST39061
12	759	75.7	871	12	BG198917	BG198917 RST18188
13	691	69.0	447	13	BG982234	BG982234 CM4-CN008
14	684	68.3	758	12	BG198969	BG198969 RST18241
15	573	57.2	594	10	AW469181	AW469181 hc78fl1.x
16	496	49.5	585	14	BQ552730	BQ552730 H4017G07-
17	492	49.1	585	10	AW469177	AW469177 hc78e12.x
18	489	48.8	376	12	BE849364	BE849364 uw88h05.y
19	489	48.8	491	10	AW913564	AW913564 uf54e02.y
20 c	484	48.3	558	9	AI799626	AI799626 to74b03.x
21 c	433.5	43.3	508	12	BF375243	BF375243 QV4-ST021
22 c	419	41.8	855	12	BG204090	BG204090 RST23484
23 c	405	40.4	307	12	BG187692	BG187692 RST6558 A
24	391.5	39.1	777	9	AU138880	AU138880
25	365	36.4	529	13	BM432955	BM432955 LJEJ7B4.a
26 c	357	35.6	829	12	BG201350	BG201350 RST20811
27 c	354	35.3	415	14	BQ084432	BQ084432 K-EST0149
28	354	35.3	482	14	BM774483	BM774483 K-EST0057
29 c	352	35.1	480	9	AI697014	AI697014 wc76h09.x
30	333	33.2	268	12	BG184955	BG184955 RST3889 A
31	327	32.6	398	14	BM773482	BM773482 K-EST0057
32	317	31.6	870	9	AJ443652	AJ443652 AJ443652
33 c	298	29.7	748	12	BE746094	BE746094 60157r3466
34	297.5	29.7	678	13	BM042049	BM042049 60361605050
35	297	29.6	956	12	BF037490	BF037490 601460605
36	296	29.5	560	10	BE683481	BE683481 183279 MA
37	296	29.5	972	12	BG174305	BG174305 60233631 2
38	293	29.2	853	14	BQ682678	BQ682678 AGENCOURT
39	293	29.2	1136	14	BQ436217	BQ436217 AGENCOURT
40	292.5	29.2	636	12	BG760177	BG760177 602733267
41	292	29.1	918	12	BG755603	BG755603 602716290
42	291.5	29.1	980	14	BG677732	BG677732 AGENCOURT
43	291	29.0	428	9	AI830044	AI830044 wj58f02.x
44	291	29.0	510	12	BF759621	BF759621 QV3-CT063
45	291	29.0	884	13	BI910512	BI910512 60306789 3

ALIGNMENTS

RESULT 1
BG184448
BG184448

LOCUS	BG184448 807 bp mRNA linear EST 21-APR-2001
DEFINITION	RST33374 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION	BG184448
VERSION	BG184448.1 GI:13706135
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 807)
AUTHORS	Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
	Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.
	, Lerner,L., Costanzo,D., McElligott,K., Boozer,S., Mays,R., Smith

, E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
,J., Danzig,J. and Ducar,M.
TITLE   Creation of genome-wide protein expression libraries using random
           activation of gene expression
JOURNAL
MEDLINE   21227151
COMMENT   Nat. Biotechnol. 19 (5), 440-445 (2001)

           Contact: Scott J. Cain
           Athersys, Inc.
           3201 Carnegie Ave, Cleveland, OH 44115, USA
           Tel: 216 431 9900
           Fax: 216 361 9596
           Email: scain@athersys.com
           High quality sequence stop: 499.
FEATURES   Location/Qualifiers
     source   1..807
                   /organism="Homo sapiens"
                   /db_xref="taxon:9606"
                   /clone_lib="Athersys RAGE Library"
                   /cell_line="HT1080"
                   /note="See 'Creation of Genome-wide Protein Expression
                   Libraries using Random Activation of Gene Expression',
                   Nature Biotechnology, in press. Note that even though the
                   cell type indicated is HT1080, since a random activation
                   method was used, these sequence tags are not necessarily
                   expressed in HT1080 under normal circumstances."
BASE COUNT   211 a   195 c   178 g   222 t   1 others
ORIGIN

US-09-895-298a-83 (1-190) x BG184448 (1-807)

Alignment Scores:
Pred. No.:
Score: 6.04e-102
                      102.00
Percent Similarity: 100.00%   Length: 807
Best Local Similarity: 100.00%   Matches: 190
Query Match: 100.00%   Conservative: 0
                      12   Mismatches: 0
                                       Indels: 0
                                       Gaps: 0

Qy   1 MetMetAsnPheGlnProProSerLysAlaTrpArgAlaSerGlnMetMetThrPhePhe 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 ATGATGAATTTCCAGCCTCCGAGCAAAAGCCTGGCGGCCTCACAGAATGAGACTTCTTC 144

Qy   21 IlePheLeuLeuPhePheProSerPheThrGlyValLeuCysThrLeuAlaIleThrIle 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   145 ATCTTCTTTGCTCTTTTTCCATCCTTCACCGGGGTCTTGTGCACCCTGGCCATCACAATC 204

Qy   41 TrpArgLeuLysProSerAlaAspCysGlyProPheArgGlyLeuProLeuPheHis 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   205 TGGAGATTGAAGCCTTCAGCTGACGTGTGGCCCTTTCGAGGTCTGCCTCTTCATTCAC 264

Qy   81 IleTyrArgAsnLeuIleGlySerValHisPhePheIleLeuThrLeuIleValLeu 100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   325 ATCTATCGGAACCTCATTGGAAGTGTGCACTCTTTTTCATCCTCACCCTCATTGTGCTA 384

Qy   101 IleIleThrTyrLeuThrTrpIleThrGlnIleThrGluLeuArgLeuLeu 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   385 ATCATCACCTACCTTACTGGCAGATCACAGAGGGAAGGAGAATTATGAGACTGCTC 444

Qy   61 SerIleTyrSerTrpIleAspThrLeuSerThrArgProGlyTyrTrpLeuTrpValValTrp 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   265 TCCATCTACAGCTGGATCGACACCCTAAGTACACGGCCCTGCCTACCTGTGGGTTGTTTGG 324

Qy   121 HisGluGlnIleIleAsnGluGlyLysAspGlyLysMetPheLeuIleGlyLeuIleLys 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   445 CATGAGCAGATCATTAATGAGGGCAAAGATAAAAAATGTTCCTGATAGAAAAATTGATCAAG 504

Qy   141 LeuGlnAspMetGluLysLysAlaAsnProSerSerLeuValLeuGluArgArgGlyVal 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   505 CTGCAGGATATGGAGAAGAAAGCAAACCCCAGCTCACTGTGTTCTGGAAAGAGAGAGGTG 564

Qy   161 GluGlnGlyPheLeuHisLeuGlyLeuGluHisSerGlySerLeuAspLeuArgSerArg 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   565 GAGCAACAAGGCTTTTTGCATTTGGGGGAACATGATGGCAGTCTTGGACTTCGGATCTAGA 624

Qy   181 ArgSerValGlnGluGlyAsnProArgAla 190
          ||||||||||||||||||||||||||||||
Db   625 AGATCAGTTCAAGAAGGTAATCCAAGGCC 654

RESULT 2
BG217799
          BG217799   777 bp   mRNA   linear
LOCUS   RST37513 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION BG217799
VERSION   BG217799.1   GI:13743820
KEYWORDS EST.
SOURCE   human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 777)
AUTHORS   Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
          Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.
          , Lerner,L., Costanzo,D., McElligott,K., Boozer,S., Mays,R., Smith
          , E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
          , J., Danzig,J. and Ducar,M.
TITLE   Creation of genome-wide protein expression libraries using random
          activation of gene expression
JOURNAL
MEDLINE   21227151
COMMENT   Nat. Biotechnol. 19 (5), 440-445 (2001)

          Contact: Scott J. Cain
          Athersys, Inc.
          3201 Carnegie Ave, Cleveland, OH 44115, USA
          Tel: 216 431 9900
          Fax: 216 361 9596
          Email: scain@athersys.com
          High quality sequence stop: 512.
FEATURES   Location/Qualifiers
     source   1..777
                   /organism="Homo sapiens"
                   /db_xref="taxon:9606"
                   /clone_lib="Athersys RAGE Library"
                   /cell_line="HT1080"
                   /note="See 'Creation of Genome-wide Protein Expression
                   Libraries using Random Activation of Gene Expression',
                   Nature Biotechnology, in press. Note that even though the
                   cell type indicated is HT1080, since a random activation
                   method was used, these sequence tags are not necessarily
                   expressed in HT1080 under normal circumstances."
BASE COUNT   196 a   193 c   173 g   214 t   1 others
ORIGIN

US-09-895-298a-83 (1-190) x BG217799 (1-777)

Alignment Scores:
Pred. No.:
Score: 9.87e-101
                      991.00
Percent Similarity: 98.95%   Length: 777
Best Local Similarity: 98.95%   Matches: 188
Query Match: 98.90%   Conservative: 0
                      12   Mismatches: 2
                                       Indels: 0
                                       Gaps: 0

Qy   1 MetMetAsnPheGlnProProSerLysAlaTrpArgAlaSerGlnMetMetThrPhePhe 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 ATGATGAATTTCCAGCCTCCGAGCAAAAGCCTGGCGGCCTCACAGAATGAGACTTCTTC 144

Qy   21 IlePheLeuLeuPhePheProSerPheThrGlyValLeuCysThrLeuAlaIleThrIle 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   145 ATCTTCTTTGCTCTTTTTCCATCCTTCACCGGGGTCTTGTGCACCCTGGCCATCACAATC 204

Qy   41 TrpArgLeuLysProSerAlaAspCysGlyProPheArgGlyLeuProLeuPheHis 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   205 TGGAGATTGAAGCCTTCAGCTGACGTGTGGCCCTTTCGAGGTCTGCCTCTTCATTCAC 264

Qy   61 SerIleTyrSerTrpIleAspThrLeuSerThrArgProGlyTyrTrpLeuTrpValValTrp 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   265 TCCATCTACAGCTGGATCGACACCCTAAGTACACGGCCCTGCCTACCTGTGGGTTGTTTGG 324

```
Qy    81 IleTyrArgAsnLeuIleGlySerValHisPhePhePheIleLeuThrLeuIleValLeu 100
         |||||||||||||||||||||||||||||||||||||||||||||||  ||||||
Db   325 ATCTATCGGAACCTCATTGGAAGTGTGCCACTTCTTTTTCATCCCTCACCCTGCGCTA 384

Qy   101 IleIleThrTyrLeuThrTyrTrpGlnIleIleThrGluGlyArgLysIleMetIleArgLeuLeu 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   385 ATCATCACCTATCTTTACTCGCAGATCATCACAGAGGGAAGGAAGATTATGATAAGGCTGCTC 384

Qy   121 HisGluGlnIleIleAsnGluGlyLysAspGlyMetProLeuLeuIleGluGlyLysLeuIleLys 140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   445 CATGAGCAGATCATTAATGAGGGCAAAGATAAAATGTTCCTGATAGAAAAATTGATCAAG 504

Qy   141 LeuGlnAspMetGluGlyLysLysAlaAsnProSerLeuValLeuGluArgArgGluVal 160
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   505 CTGCAGGATATGGAGAAGAAAGCAAACCCCAGCTCACTTGTTCTGGAAAGGAGAGAGGTG 564

Qy   161 GluGlnIleGlyPheLeuHisLeuGlyLeuHisAspGlySerLeuAspLeuArgSerArg 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   565 GAGCAAACAAGGCTTTTTGCATTTGGGGGAACATGATGGCAGTCTTGACTTGCGATCTANA 624

Qy   181 ArgSerValGlnGlyGlyAsnProArgAla 190
         ||||||||||||||||||||||||||||
Db   625 AGATCAGTTCAAGAAGGTAATCCAAGGGCC 654
```

RESULT 3

```
LOCUS       AW582253                 691 bp    mRNA    linear    EST 16-MAR-2000
DEFINITION  QV4-ST0212-120100-f09 ST0212 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW582253
VERSION     AW582253.1  GI:7257302
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 691)
  AUTHORS   HCGP http://www.ludwig.org.br/ORESTES.
  TITLE     The FAPESP/LICR Human Cancer Genome Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL.
            (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=QV4&t2=QV4-ST0212-
            120100-075-f09&t3=2000-01-12&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 5
            High quality sequence stop: 667.
FEATURES             Location/Qualifiers
     source          1..691
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="ST0212"
                     /dev_stage="Adult"
                     /note="Organ: stomach; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from ORESTES PCR (U.S. Letters Patent Application
                     No. 196,716 - Ludwig Institute for Cancer Research)
                     profiles into the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT      179 a    167 c    157 g    188 t
ORIGIN
```

```
Score:                 970.50              Matches:         189
Percent Similarity:     99.47%             Conservative:      0
Best Local Similarity:  99.47%             Mismatches:        1
Query Match:            96.86%             Indels:            1
DB:                     10                 Gaps:              0
```

US-09-895-298A-83 (1-190) x AW582253 (1-691)

```
Qy    1 MetMetAsnPheGlnProProSerLysAlaThrpArgArgAlaSerGlnMetMetThrPhePhe 20
        ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   28 ATGATGAATTTCCAGCCTCCGAGCA--GCCTGGCGGCGGCCTCACAGATGATGACTTTCTTC 85

Qy   21 IlePheLeuLeuPhePheProSerPheThrGlyValLeuCysThrLeuAlaIleThrIle 40
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   86 ATCTTCTTGCTCTTTTTCCCATCCTTCACCGGGGTCTTGTGCACCCTGGCCATCACCATC 145

Qy   41 TrpArgLeuGlyProSerAlaAspCysGlyProPheArgGlyLeuProLeuPheIleHis 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  146 TGGAGATTGGAAGCCTTCAGCGACTGTGGCCCCTTTTCGAGGTGCCGCTCTTCATTCAC 205

Qy   61 SerIleTyrSerTrpIleAspThrLeuSerThrArgProGlyTyrLeuTrpValTrp 80
        |||||||||||||||||||||||||||||||||||||||||||||||
Db  206 TCCAATCTACAGCCTGGATCGACACCCTAAGTACACGGCCCTGGCTACCCTGTCGGGTG 265

Qy   81 IleTyrArgAsnLeuIleGlySerValHisPhePhePheIleLeuThrLeuIleValLeu 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  266 ATCTATCGGAACCTCATTGGAAGTGTGCCACTTCTTTTTCATCCCTCACCCTGCGCTA 325

Qy  101 IleIleThrTyrLeuThrTyrTrpGlnIleIleThrGluGlyArgLysIleMetIleArgLeuLeu 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  326 ATCATCACCTATCTTTACTCGCAGATCATCACAGAGGGAAGGAAGATTATGATAAGGCTGCTC 385

Qy  121 HisGluGlnIleIleAsnGluGlyLysAspGlyMetProLeuLeuIleGluGlyLysLeuIleLys 140
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  386 CATGAGCAGATCATTAATGAGGGCAAAGATAAAATGTTCCTGATAGAAAAATTGATCAAG 445

Qy  141 LeuGlnAspMetGluGlyLysLysAlaAsnProSerLeuValLeuGluArgArgGluVal 160
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  446 CTGCAGGATATGGAGAAGAAAGCAAACCCCAGCTCACTTGTTCTGGAAAGGAGAGAGGTG 505

Qy  161 GluGlnIleGlyPheLeuHisLeuGlyLeuHisAspGlySerLeuAspLeuArgSerArg 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  506 GAGCAAACAAGGCTTTTTGCATTTGGGGGAACATGATGGCAGTCTTGACTTGCGATCTAGA 565

Qy  181 ArgSerValGlnGlyGlyAsnProArgAla 190
        ||||||||||||||||||||||||||||
Db  566 AGATCAGTTCAAGAAGGTAATCCAAGGGCC 595
```

RESULT 4

```
LOCUS       BG194960                 768 bp    mRNA    linear    EST 21-APR-2001
DEFINITION  RST14018 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG194960
VERSION     BG194960.1  GI:13716535
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 768)
  AUTHORS   Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
            Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.
            , Lerner,L., Costanzo,D., McElligott,K., Boozer,S., Mays,R., Smith
            ,J., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
            ,J., Danzig,J. and Ducar,M.
  TITLE     Creation of genome-wide protein expression libraries using random
            activation of gene expression
  JOURNAL   Nat. Biotechnol. 19 (5), 440-445 (2001)
  MEDLINE   21227151
COMMENT     Contact: Scott J. Cain
            Athersys, Inc.
            3201 Carnegie Ave, Cleveland, OH 44115, USA
            Tel: 216 431 9900
```

Alignment Scores:
Pred. No.:

Alignment Scores: 1.71e-98 Length: 691

Left column

```
                  Fax: 216 361 9996
                  Email: scain@athersys.com
                  High quality sequence stop: 469.

FEATURES          Location/Qualifiers
     source       1..768
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone_lib="Athersys RAGE Library"
                  /cell_line="HT1080"
                  /note="See 'Creation of Genome-wide Protein Expression
                  Libraries using Random Activation of Gene Expression',
                  Nature Biotechnology, in press. Note that even though the
                  cell type indicated is HT1080, since a random activation
                  method was used, these sequence tags are not necessarily
                  expressed in HT1080 under normal circumstances."

BASE COUNT     197 a    188 c    171 g    212 t
ORIGIN

US-09-895-298A-83 (1-190) x BG194960 (1-768)

Alignment Scores:
Pred. No.:
Score:                 957.00          Length:        768
Percent Similarity:    97.89%          Matches:       185
Best Local Similarity: 97.37%          Conservative:    1
Query Match:           95.51%          Mismatches:      4
DB:                                    Indels:          1
                                       Gaps:            0

6.26e-97

Qy    1 MetMetAsnPheGlnProProSerLysAlaThrArgAlaSerGlnMetMetThrPhePhe  20
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 ATGATGAATTTCCAGCCTCCGAGCAAAGCCACTCGGGCCTCCCAGATGATGACTTCTTC 144

Qy   21 IlePheLeuLeuPhePheProSerPheThrGlyValLeuCysThrLeuAlaIleThrIle  40
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db  145 ATCTTCTTGCTCTTTTTCCCATCCTTCACCGGGGTCTTGTGCACCCTGGCCATCACCATC 204

Qy   41 TrpArgLeuLysProSerAlaAspCysGlyProLeuProLeuPheIleHis  60
        |||||||||||||||||||||||||||||||||||||||||||||
Db  205 TGGAGATTGAAGCCTTCAGCTGACTGTGGCCCTTTTCGAGGTCCTGCCTTCTTCATCCAC 264

Qy   61 SerIleTyrSerTrpIleLeuSerThrArgProGlyTyrTyrLeuTrpValValTrp  80
        ||||||||||||||||||||||||||||||||||||||||||
Db  265 TCCAATCTACAGCTGGATCGACACACCCTAAGTCACGGGCCCTGGTTACTGTGGGTGGTTGG 324

Qy   81 IleTyrArgAsnLeuIleLeuGlySerValHisPhePheIleLeuThrLeuIleValLeu 100
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db  325 ATCTATCGGAACCTCATCCTGGGAAGTGTGCACTTCTTTATCATCCTCACCCTCAATGCCTA 384

Qy  101 IleIleThrThrTyrLeuTyrTrpGlnIleThrGlyIleArgLysIleMetIleArgLeuLeu 120
        |||||||||||||||||||||||||||||||||||||||||||||
Db  385 ATCATCACCTACCTTTACTGGCAGATCACAGAGGGAAGGAAGATTATGATAAGGCTGCTC 444

Qy  121 HisGlnGlnIleIleIleAsnGluGlyLysMetPheLeuIleGluGlyLysIleLeuLys 140
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db  445 CATGAGCAGATCATTAATGAGGGCCAAAGATAAAAATGTTCCTGATAGGAAAAATTGATCAAG 504

Qy  141 LeuGlnAspMetGlyLysLysAlaAsnProSerSerLeuValLeuGlyArgArgGluVal 160
        ||||||||||||||||||||||||||||||||||||||||||||
Db  505 CTGCAGGATATGGGAAGAAAGCCAACCCCAGCCTCACTTGGTCTGGTGAAAGAGAGAGTG 564

Qy  161 GluGlnGlnGlyPheLeuGluHisLeuGlyGluHisAspGlySerLeuAspSerArg 180
        |||||||||||||||||||||||||||||||||||||||
Db  565 GAGCAACAAGGCTTTTGCATTGGGGAACATGATGGGCAGTTTGGATTCAA 624

Qy  181 ArgSerValGlnGlyGlyAsnProArgAla 190
        ||||:|||||||||||||||||:::
Db  625 AGACA-GTTCAAGAAGGTAATCCAAGGGCC 653
```

Right column

```
FEATURES          Location/Qualifiers
     source       1..789
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone_lib="Athersys RAGE Library"
                  /cell_line="HT1080"
                  /note="See 'Creation of Genome-wide Protein Expression
                  Libraries using Random Activation of Gene Expression',
                  Nature Biotechnology, in press. Note that even though the
                  cell type indicated is HT1080, since a random activation
                  method was used, these sequence tags are not necessarily
                  expressed in HT1080 under normal circumstances."

BASE COUNT     191 a    194 c    171 g    226 t    7 others
ORIGIN

US-09-895-298A-83 (1-190) x BG221275 (1-789)

Alignment Scores:
Pred. No.:
Score:                 923.00          Length:        789
Percent Similarity:    96.34%          Matches:       183
Best Local Similarity: 95.81%          Conservative:    1
Query Match:           92.12%          Mismatches:      6
DB:                                    Indels:          2
                                       Gaps:            0

4.15e-93

Qy    1 MetMetAsnPheGlnProProSerLysAlaThrArgAlaSerGlnMetMetThrPhePhe  20
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  188 ATGATGAATTTCCAGCCTCCGAGCTAAGCCTGGGGCCTCCCAGATGATGACTTCTTC 247

Qy   21 IlePheLeuLeuPheProSerPheThrGlyValLeuCysThrLeuAlaIleThrIle  40
        ||||||||||||||||||||||||||||||||||||||||||||
Db  248 ATCTTCTTGCTCTTTTTCCCATCCTTCACCGGGGTCTTGTGCACCCTGGCCATCACCATC 307

Qy   41 TrpArgLeuLysProSerAlaAspCysGlyProLeuProLeuPheIleHis  60
        ||||||||||||||||||||||||||||||||||||||||||||
Db  308 TGGAGATTGAAGCCTTCAGCTGACTGTGGCCCTTTTCGAGGTCCTGCCTTCTTCATCCAC 367

Qy   61 SerIleTyrSerTrpIleLeuSerThrArgProGlyTyrTyrLeuTrpValValTrp  80
        ||||||||||||||||||||||||||||||||||||||||||||
Db  368 TCCAATCTACAGCTGGATCGACACCTAAGTACACGGGCCCTGGTTACTGTGGGTGGTTGG 427

Qy   81 IleTyrArgAsnLeuIleLeuGlySerValHisPhePheIleLeuThrLeuIleValLeu 100
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db  428 ATCTATCGGAACCTCATCCTGGGAAGTGTGCACTTCTTTATCATCCTCACCCTGTTGGG 487

Qy  101 IleIleThrThrTyrLeuTyrTrpGlnIleThrGlyIleArgLysIleMetIleArgLeuLeu 120
        ||||||||||||||||||||||||||||||||||||||||||||
Db  488 ATCATCACCTATCTTTACTGGCAGATCACAGAGGGAAGGAAGATTATGATAAGGCTGCTC 547

Qy  121 HisGlnGlnIleIleIleAsnGluGlyLysMetPheLeuIleGluGlyLysIleLeuLys 140
        |||||||||||||||||||||||||||||||||||||||||||||||
Db  ...
```

Bottom (left column)

```
RESULT 5
LOCUS       BG221275     789 bp    mRNA    linear    EST 21-APR-2001
DEFINITION  RST41079 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG221275
```

RESULT 6

```
Db  548 CATGAGCAGATCATTAATGAGGGCAAAGATAAAATGTTCCTGATAGAGAAAATTGATCAGG 607
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||::::
Qy  727 AGATCAGTTCAAGAAGGTAAAATCCCAGGGCC 757

Qy  181 ArgSerValGlnGlyLys-AsnProArgAla 190
Db  668 GAGNCACAAGGCTTTTTGCATTGNG-GAACANTGGACGCATGTTGCGATCTGGA 726

Qy  161 GluGlnGlyPheLeuHisLeuGlyLySerLeuAspLeuArgSerArg 180
        |||||||||||||||||||||||||||||||||||||||||||||||
Db  608 CTGCCAGGAFAATGGAGAAGAAANACCCCACGTCACTTGTTCTTCGGANAGGAGAGNAGTG 667

Qy  141 LeuGlnAsnProMetGluLysLysAlaAsnProSerSerLeuValLeuGluArgGlyVal 160
Db  181 ArgSerValGlnGlyLys-AsnProArgAla 190
```

LOCUS   BG216071
DEFINITION RST35885 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG216071   837 bp   mRNA   linear   EST 21-APR-2001
VERSION   BG216071.1 GI:13742220
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 837)
  AUTHORS   Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
            Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.
            , Lerner,L., Costanzo,D., McElligott,K., Boozer,S., Mays,R., Smith
            , E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
            , J., Danzig,J. and Ducar,M.
  TITLE   Creation of genome-wide protein expression libraries using random
            activation of gene expression
  JOURNAL   Nat. Biotechnol. 19 (5), 440-445 (2001)
  MEDLINE   21227151
COMMENT   Contact: Scott J. Cain
            Athersys, Inc.
            3201 Carnegie Ave, Cleveland, OH 44115, USA
            Tel: 216 431 9900
            Fax: 216 361 9596
            Email: scain@athersys.com
            /clone_lib="Athersys RAGE Library"
            /note="See 'Creation of Genome-wide Protein Expression
            Libraries using Random Activation of Gene Expression',
            J., Danzig,J. and Ducar,M.
            Nature Biotechnology, in press. Note that even though the
            cell type indicated is HT1080, since a random activation
            method was used, these sequence tags are not necessarily
            expressed in HT1080 under normal circumstances."
FEATURES   Location/Qualifiers
     source   1..837
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
            High quality sequence stop: 537.

ORIGIN

BASE COUNT   208 a   209 c   184 g   235 t   1 others

US-09-895-298a-83 (1-190) x BG216071 (1-837)

Alignment Scores:
Pred. No.:
Score: 4.51e-92
                        914.00
Percent Similarity: 96.84%
Best Local Similarity: 95.79%
Query Match: 91.22%
DB: 12

Length: 837
Matches: 182
Conservative: 2
Mismatches: 6
Indels: 2
Gaps: 0

```
Qy   1 MetMetAsnPheGlnProProSerLysAlaTrpArgAlaSerGlnMetMetThrPhePhe 20
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124 ATGATGAATTTCCACCTCCCAGCAAAGCCTGGCGGGCCTCACAGATGATGACTTTCTTC 183

Qy  21 IlePheLeuLeuPhePheProSerPheThrGlyValLeuCysThrLeuAlaIleThrIle 40
```

RESULT 7

```
Db  184 ATCTTCTTTGCTCTCTTTTTTCCTTCCATCCTTCACCCGGGGTCTTGTGTGCACCCATC 243

Qy   41 TrpArgLeuLysProSerAlaAspCysGlyProPheArgGlyLeuProLeuPheIleHis 60
Db  244 TGGAGATTGAAGCCTTCAGCTGACTGTGCCCTTTTCGGAGTGTCCCTCTTCATTCAC 303

Qy   61 SerIleTyrSerTrpIleAspThrLeuSerThrArgProGlyTyrTyrLeuTrpValTrp 80
Db  304 TCCATCTACAGCTGGATCGACACCCTAAGTACACGGCCTGGGTACTATCTGTGGGTTTGG 363

Qy   81 ArgSerValGlnGlyLysHisProPhePheLeuArgLeuIleLeuValLeu 100
Db  364 AGCTATCGGAACCTCATTGGAAGTGTGCACTTCTTTCATCCTCACCCTCACCCTCATTGTGCTA 423

Qy  101 IleIleThrTyrTrpLeuTyrTrpGlnIleIleThrGluProLysLysLysIleMetIleArgGluLeu 120
Db  424 ATCATCACCTATCTTCTTACTCGGCCAGATCACAGATGAAGGAAGAAGATTATGATAAGGCTGCTC 483

Qy  121 HisGlyGlnIleIleAlaAsnGluValLysLysAspLeuMetPheLeuGluLysLysLeuIleLys 140
Db  484 CATGAGCAGATCATTAATGAGGGCAAAGATAAAATGTTCCTGATAGAAAATTGATCAAG 543

Qy  141 LeuGlnAsnProMetGluLysLysAlaAsnProSerSerLeuValLeuGluArgGlyVal 160
Db  544 CTGCCAGGATATGGAGAAGAAGACAAACCCCACGTCACTTGTTCTTCGGAAAGGAGAGAGGTG 603

Qy  161 GluGlnGlyPheLeuHisLeuGlyLySerLeuAspLeuArgSerArg 180
Db  604 GAGCA-CAAGGCTTTTTTGCATTTGGGGGAACANTGAGGGCAG-CTTGACTTGGGGGATCTAGA 661

Qy  181 ArgSerValGlnGlyLysAsnProArgAla 190
Db  662 AGATCACGTCAAGAAGGTGANCCAAGGGCC 691
```

LOCUS   BG183902
DEFINITION RST2815 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG183902   823 bp   mRNA   linear   EST 21-APR-2001
VERSION   BG183902.1 GI:13705589
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 823)
  AUTHORS   Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
            Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.
            , Lerner,L., Costanzo,D., McElligott,K., Boozer,S., Mays,R., Smith
            , E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
            , J., Danzig,J. and Ducar,M.
  TITLE   Creation of genome-wide protein expression libraries using random
            activation of gene expression
  JOURNAL   Nat. Biotechnol. 19 (5), 440-445 (2001)
  MEDLINE   21227151
COMMENT   Contact: Scott J. Cain
            Athersys, Inc.
            3201 Carnegie Ave, Cleveland, OH 44115, USA
            Tel: 216 431 9900
            Fax: 216 361 9596
            Email: scain@athersys.com
            /clone_lib="Athersys RAGE Library"
            /note="See 'Creation of Genome-wide Protein Expression
            Libraries using Random Activation of Gene Expression',
            J., Danzig,J. and Ducar,M.
            Nature Biotechnology, in press. Note that even though the
            cell type indicated is HT1080, since a random activation
FEATURES   Location/Qualifiers
     source   1..823
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
            High quality sequence stop: 449.

BASE COUNT   211 a   203 c   180 g   229 t

ORIGIN

US-09-895-298A-83 (1-190) x BG183902 (1-823)

Alignment Scores:
Pred. No.:
Score: 3.48e-91
                          906.00
Percent Similarity: 96.34%
Best Local Similarity: 95.81%
Query Match: 90.42%
DB: 12

Length: 823
Matches: 183
Conservative: 1
Mismatches: 6
Indels: 3
Gaps: 0

Qy   1 MetMetAsnPheGlnProProSerLysAlaTrpArgAlaSerGlnMetMetThrPhePhe 20
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   109 ATGATGAATTTCCAGCCTCCGAGCAAAGCCTGGCGGGCCTCACAGATGATGACTTCTTC 168

Qy   21 IlePheLeuLeuPhePheProSerPheThrGlyValLeuCysThrLeuAlaIleThrIle 40
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   169 ATCTTCTTTGCTTTTTCCATCCTTCACCGGGGTCTTGTGCACCCTGGCCATCACCATC 228

Qy   41 TrpArgLeuLysProSerAlaAspCysGlyProPheArgGlyLeuProLeuPheIleHis 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   229 TGGAGATTGAAGCCTTCAGCTGACTGTGGCCCTTTTCGAGGTCGTTGCCTCTTCATTCAC 288

Qy   61 SerIleTyrSerTrpIleSerThrLeuSerThrArgProGlyTyrLeuTrpValValTrp 80
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   289 TCCATCTACAGCTGGATCAGCACACCCTAAGTACACGGCCTACCCTGTGGTTGTTTGG 348

Qy   81 IleTyrArgAsnLeuIleGlySerValHisPhePhePheIleLeuThrLeuIleValLeu 100
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   349 ATCTATCGGAACCTCATTGGAAGTGTGCACTTCTTTTTCATCCTCACCCTCATTGTGCTA 408

Qy   101 IleIleThrTyrLeuTyrTrpGlnIleIleThrGluGlyArgLysIleMetIleLeuArgLeuLeu 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   409 ATCATCACCTATCTTTACTGGCAGATCACAGAGGGAAGGAAGATTATGATAAGGCTGCTC 468

Qy   121 HisGlyGlnIleIleIleAsnGluGlyLysMetPheLeuIleGluGlyLysIleLeuLys 140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   469 CATGAGCAGATCATCTATGAGGGCAAAGATAAAAGTTCCTGATAGAGAAAATTGATCAAG 528

Qy   141 LeuGlnAsnMetGlyLysLysAlaAsnProSerSerLeuValLeuGluArgArgGluVal 160
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   529 CTGCAGGATAPCGGAGAAAAATACCAAACCCCAGCTCACTTGTTCTCGGAAAGAGAGAGGG 588

Qy   161 GluGlnGlyPheLeuHisLeuGlyTyrSerLeuAspSerArgSerArg 180
           ||| ||||||||||||||||||||||||||||||||||||||||::::
Db   589 GAAGCAACAAGGCTTTTTGCATTGGGGGAACATGATGGCAGTCTTGACTT-CGATCTAA 647

Qy   180 gArgSerValGlnGluGlyAsnProArgAla 190
           ::|||||||||||||||||||||||
Db   648 AATATCAGTTCAAGAGAGGGAT-CCAAGGGCC 677

RESULT 8

LOCUS   BG217267   810 bp   mRNA   linear   EST 21-APR-2001
DEFINITION   RST36999 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG217267
VERSION   BG217267.1   GI:13743288
KEYWORDS   EST.
SOURCE   human.
  ORGANISM   Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 810)
  AUTHORS   Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
               Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.
               , Lerner,L., Costanzo,D., McElligott,K., Boozer,S., Mays,R., Smith
               ,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
               ,J., Danzig,J. and Ducar,M.
  TITLE   Creation of genome-wide protein expression libraries using random

---

COMMENT   method was used, these sequence tags are not necessarily
               expressed in HT1080 under normal circumstances."

JOURNAL   Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE   21227151

COMMENT   activation of gene expression

Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@athersys.com
High quality sequence stop: 466.

FEATURES   Location/Qualifiers
     source   1..810
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="Athersys RAGE Library"
                     /cell_line="HT1080"
                     /note="See 'Creation of Genome-wide Protein Expression
                     Libraries using Random Activation of Gene Expression',
                     Nature Biotechnology, in press. Note that even though the
                     cell type indicated is HT1080, since a random activation
                     method was used, these sequence tags are not necessarily
                     expressed in HT1080 under normal circumstances."

BASE COUNT   207 a   196 c   180 g   227 t

ORIGIN

US-09-895-298A-83 (1-190) x BG217267 (1-810)

Alignment Scores:
Pred. No.:
Score: 9.6e-91
                          902.00
Percent Similarity: 95.26%
Best Local Similarity: 93.16%
Query Match: 90.02%
DB: 12

Length: 810
Matches: 177
Conservative: 4
Mismatches: 7
Indels: 2
Gaps: 0

Qy   1 MetMetAsnPheGlnProProSerLysAlaTrpArgAlaSerGlnMetMetThrPhePhe 20
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 ATGATGAATTTCCAGCCTCCGAGCAAAGCCTGGCGGGCCTCACAGATGATGACTTTCTTC 144

Qy   21 IlePheLeuLeuPhePheProSerPheThrGlyValLeuCysThrLeuAlaIleThrIle 40
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   145 ATCTTCTTTGCTCTTTTTTCCATCCTTCACCGGGGTCTTGTGCACCCTGGCCATCACCATC 204

Qy   41 TrpArgLeuLysProSerAlaAspCysGlyProPheArgGlyLeuProLeuPheIleHis 60
           |||||||||||||||||||||||||||::|||||||||||||||||||||||||||||||
Db   205 TGGAGATTGAAGCCTTCAGCCGCTGACGTCGGCCCTTTTCGAGGTCGTTGCCTCTTCATTCAC 264

Qy   61 SerIleTyrSerTrpIleSerThrLeuSerThrArgProGlyTyrLeuTrpValValTrp 80
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   265 TCCATCTACAGCTGGATCAGCACACCCTAAGTACACGGCCTACCCTGTGGTTGTTTGG 324

Qy   81 IleTyrArgAsnLeuIleGlySerValHisPhePhePheIleLeuThrLeuIleValLeu 100
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   325 ATCTATCGGAACCTCATTGGAAGTGTGCACTTCTTTTTCATCCTCACCCTCATTGTGGCTA 384

Qy   101 IleIleThrTyrLeuTyrTrpGlnIleIleThrGluGlyArgLysIleMetIleLeuArgLeuLeu 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   385 ATCATCACCTATCTTTACTGGCAGATCACAGAGGGAAGGAAGATTATGATAAGGCTGCTC 444

Qy   121 HisGlyGlnIleIleIleAsnGluGlyLysMetPheLeuIleGluGlyLysIleLeuLys 140
           |||||||||||| ||| ||||||||||||||||||||||||||||||||||
Db   445 CATGAGCAGATCATTAATGAGGGCAAAGATAAAAAGTTGGTCCTGATAGAGAAAATTGATCAAG 504

Qy   141 LeuGlnAsnMetGlyLysLysAlaAsnProSerSerLeuValLeuGluArgArgGluVal 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   505 CTGCAGGATATGGGACGAAAGCCAAACCCCAGCTCACTTGTTCTCGGAAAGAGAGAGGTG 564

Qy   161 GluGln-GlnGlyPheLeuHisLeu-GlyGluHisAspLeuGlySerLeuAspLeuArgSerA 180
           ||||| ||||||||||||| ||| |||||||||||||||||||||||||||||||||||
Db   565 GAGCAACCAGGCTTTTTGATTTTGAGGGAACATGATGGGCAGTCTTGACTTGCGATCTA 624

Qy   180 rgArgSerValGlnGluGlyAsnPro 188
           |||::|||||::||::::|||||
Db   625

```
Db      625  GAAAATCAATTCAGAAGGAATCCCA  650

RESULT 9

LOCUS       AK016573                2721 bp    mRNA    linear   HTC 19-JAN-2002
DEFINITION  Mus musculus adult male testis cDNA, RIKEN full-length enriched
            library, clone:4932443L08:homolog to CDNA FLJ13593 FIS, CLONE
            PLACE1009493, full insert sequence.
ACCESSION   AK016573
VERSION     AK016573.1  GI:12855380
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA,
            clone_lib:RIKEN full-length enriched mouse cDNA library
            clone:4932443L08.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20493374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
            Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
            Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
            Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
            Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
            Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
            Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
            Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
            Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
            Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
            Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
            Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
            Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
            Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
            Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
            Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
            and Hayashizaki,Y.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409 (6821), 685-690 (2001)
  MEDLINE   21085660
   PUBMED   11217851
REFERENCE   5  (bases 1 to 2721)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
            Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C.,
            Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T.,
            Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F.,
            Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,
            Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,
            Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,
            Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,
            Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and
            Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
            Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
COMMENT     Direct Submission
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in Riken.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. cDNA went
            through one round of normalization to Rot = 10.0 and subtraction to
            Rot = 100.0. Second strand cDNA was prepared with the primer
            adapter of sequence [5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAAATCACCCCCCCCCCCCCC 3']. cDNA was cleaved
            with BamHI and XhoI. cDNA of size comprised between 3 and 7 kb was
            selected before cloning. Vector: a modified pBluescript KS(+) after
            bulk excision from Lambda FLC I. Cloning sites, 5' end: Sall; 3'
            end: BamHI. Host: DH10B.
FEATURES             Location/Qualifiers
     source          1..2721
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="4932443L08"
                     /sex="male"
                     /tissue_type="testis"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     CDS             207..2480
                     /note="data source:SPTR, source key:Q9H817, evidence:ISS
                     homolog to CDNA FLJ13593 FIS, CLONE PLACE1009493
                     putative"
                     /codon_start=1
                     /protein_id="BAB30314.1"
                     /db_xref="GI:12855381"
                     /db_xref="MGD:MGI:1921674"
                     /translation="MQSDQVDEIIIEVENVPSGVQNHLVSSQIALRKSSANPAFCVL
                     SSSAADRVDCQIFNPGNDRHRHRLLRFSSLNESISQIYHGSECLVTDESCTFHETVQG
                     QKLIASIPMTTRDRIKITRNQPRTMQEKRELRKIVDKERNKQSHGILEANCAQCLG
                     SLSLTTRTRNLSELINYITLWQKRFKVLSYFSVLSYSFSFSRMLLKFNIFSYN
                     NFSFIIPQFTVGAKNTLQFTGLEFFTGAGYFGDVMYYGIYGYTNSTIRHRMGGASYNM
                     QLAYIFIGACLVVCFFSLLFSMAYFRNNFINPHIYSRGIAKLIFCNDFTVTHEKAV
                     KLKQKNLSTEIERENLSELRQENYRLTFNQOLTRFSAHVAANLVSTGVAVAACAYAYYL
                     AEYNSEFLKTHRNPGAVLLIPFVVSCINLAVPRFYSMFRLVERYEIPXQEEVYLLVRN
                     IFLKISIVGILCYWLNIVALSGEECWETILGQDIYRLLLMDFVFSLADSLLGEFLRR
                     LIGMKFTRSLSLQEFDIARNVLELIYAQTLTWLGIFECPLLPFIQMTIFMYVKNVS
                     LMMNFQPPSKAWRASQMTFMFIFLFFPFSFTGVLCTLAITTWRLKPSADCGPPRGLPS
                     FIQSISWIDTLSRPGYLWVVWIYQNLIQSVHFFILTLLIVLIITYLYWQITEGRKV
                     MIRLIHEQIINEGKDKMFLIEKLTKLIQDMEKRVNPSALDLERREVEPQIPLIHLEELGA
                     APDLRLRRSAQEENPIA"
     polyA_signal    2702..2707
                     /note="putative"
     polyA_site      2721
```

```
BASE COUNT      678 a    692 c    644 g    707 t
ORIGIN          /note="putative"

Qy

US-09-895-298A-83 (1-190) x AK016573 (1-2721)

Alignment Scores:
Pred. No.:           4.08e-85
Score:               857.00                  Length:                2721
Percent Similarity:  90.53%                  Matches:               166
Best Local Similarity: 87.37%                Conservative:          6
Query Match:         85.53%                  Mismatches:            18
                     11                      Indels:                0
                                             Gaps:                  0

Qy     1 MetMetAsnPheGlnProProSerLysAlaThrProArgAlaSerGlnMetMetThrProPhe 20
Db  1908 ATGATGAAACTTCCAGCCCCAAGCAAGGCCCTGGCGGCGTCGGCGGCTCACAGATGATGACTTTCTTC 1967

Qy    21 IleProPheLeuPheProSerPheThrGlyValValLeuCysThrLeuAlaIleThrIle 40
Db  1968 ATTTTCTTTGCTCTTCTTTCCCATCTTTCACGGGAGTGCTGTGCCACCCTGGCTATCACCATC 2027

Qy    41 TrpArgArgLeuGlyProSerAlaAspCysGlyGlyProPheArgGlyGlyLeuProLeuPheIleHis 60
Db  2028 TGGAGGACTGGAAACCTTCGGCTGACTGTGGCGGTTCCGAGGTGTGCCTTCCCTTCCTCATCCAA 2087

Qy    61 SerIleTyrSerTrpLeuSerThrArgProGlyTyrLeuTrpValValTrp 80
Db  2088 TCCATCTACAGCTGGATTGACACTCTGACGCCGCAGGCCTGCCTACCCTGTGGGTCCGTCTGG 2147

Qy    81 IleTyrArgAsnLeuIleGlySerValHisPhePhePheIleLeuThrLeuIleValLeu 100
Db  2148 ATCTACACCAGAAATCTCATTGGAAGTGTGCACTTCTTCTTCATCCTCACCCTCATTGTGCTA 2207

Qy   101 IleIleLeuThrTyrLeuTyrTrpGlnIleThrGlyValArgLysIleMetIleArgLeuLeu 120
Db  2208 ATCACCACCTATCTTTACTGGCAGATCACAGGAGGGTCAAGAAGATCATGATTCGATTGCTC 2267

Qy   121 HisGlyGlnIleLeuAsnGlyGlyLysAspGlySerLeuAspLeuArgSerArg 140
Db  2268 CATGAACAGATCATTAATGAGGGCAAAGATAAGAGTTCCTGATAGAGAAAATTGACCAAA 2327

Qy   141 LeuGlnAsnAspMetGluLysLysAlaAsnProSerSerLeuValLeuGluArgArgGluVal 160
Db  2328 CTACAGAATATGGAGAAGAAAGCTAACCCCAGTGCCCTTGACCTGGAAAGGAGAGAGGTG 2387

Qy   161 GluGlnGlnGlyPheLeuHisLeuGlyLysAspGlySerLeuAsnLeuArgSerArg 180
Db  2388 GAGCCACAAAATTCCTTTGCATTTGGAGGAACTTGGTGCTGCTCCTCGACCTGCGATTGAGG 2447

Qy   181 ArgSerValGlnGluGlyAsnProArgAla 190
Db  2448 CGGTCAGCACACAAGAGAGGAGAAATCCCATAGCC 2477
```

TITLE   Shotgun sequencing of the human transcriptome with ORF expressed
              sequence tags
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE   20202663
COMMENT   Contact: Simpson A.J.G.
              Laboratory of Cancer Genetics
              Ludwig Institute for Cancer Research
              Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
              Brazil
              Tel: +55-11-2704922
              Fax: +55-11-2707001
              Email: asimpson@ludwig.org.br
              This sequence was derived from the FAPESP/LICR Human Cancer Genome
              Project. This entry can be seen in the following URL
              (http://www.ludwig.org.br/scripts/getthtml2.pl?ti=QV4-ST0212-
              261199-045-c09&t3=1999-11-26&t4=1)
              Seq primer: puc 18 forward
              High quality sequence start: 14
              High quality sequence stop: 36.
FEATURES   Location/Qualifiers
     source   1..698
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="ST0212"
                     /note="Organ: stomach; Vector: pucl8; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 - Ludwig Institute for Cancer Research)
                     profiles into the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
                     /dev_stage="Adult"

```
BASE COUNT      177 a    173 c    164 g    184 t
ORIGIN

US-09-895-298A-83 (1-190) x BF375244 (1-698)

Alignment Scores:
Pred. No.:           2.46e-79
Score:               799.50                  Length:                698
Percent Similarity:  90.22%                  Matches:               163
Best Local Similarity: 88.59%                Conservative:          3
Query Match:         79.79%                  Mismatches:            16
                     12                      Indels:                3
                                             Gaps:                  1

Qy     9 LysAlaThrProArgAlaSerGlnMet--MetThrPhePheLeuPheLeuPhePr 27
Db     1 CAAGCCTGGCGGCGGCCTAACAGACTGATAGACTTTCTATCATCTTCTTGCTCTPATTTCCC 87

Qy    27 oSerPheThrGlyValLeuCysThrLeuAlaIleThrIleTrpArgArgLeuGlyProSerAl 47
Db   115 ATCCTTCACCGGGGTCTTGTGCACCCTGGCCATCACCATCTGGAGATTGAAGCCTTCACC 174

Qy    47 aAspCysGlyProPheArgGlyGlyLeuProLeuPheIleHisSerIleTyrSerTrpLeuAs 67
Db   175 TGACTGTGGCCCCTTTTTGCAGGTGTGCCTCTTCCTCATTCACTCCATCTACAGCTGGATCGA 234

Qy    67 pThrLeuSerThrArgProGlyTyrLeuTrpValValTrpIleTyrArgAsnLeuIleGl 87
Db   235 CACCCTAAGTACACGCCGGCCTGGCTACCCTGTGGGTTGTTTGGATCTATCCGGAACCTCATTGG 294

Qy    87 ySerValHisPhePhePheIleLeuThrLeuIleValLeuIleIleLeuThrTyrLeuTyr 107
Db   295 AAGTGTGCACTTCTTTTTCATCCTCACCCTCATTGTGCTAATCATCACCTACCTTTACTG 354

Qy   107 pGlnIleThrGlyValArgLysIleMetIleArgLeuLeuHisGlyGlnIleLeuAsnGl 127
Db   355 GCAGATCACAGAGGGAGGAAAGAAGATCATGATTCAGGCTGCTCCATGACCAGAGATCATTAATGA 414

Qy   127 uGlyGlyLysAspGlySerMetPheLeuuHisLeuGlyLysLeuGlnAspMetGluLysSLy 147
Db   415 GGGCAACGATCAAAATGTTCCTGATAGAAACATTGATCAAGCTGCCAGGATGGCGACGAC 474

Qy   147 sAlaAsnProSerSerLeuValLeuGluArgArgGluValGlnGlnGlyPheLeuHl 167
```

REFERENCE   1 (bases 1 to 698)
AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
              Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
              Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
              Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
              ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
              Simpson,A.J.
ORGANISM   Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
SOURCE   human.
KEYWORDS   EST.
VERSION   BF375244.1 GI:11337269
ACCESSION   BF375244
DEFINITION QV4-ST0212-261199-045-c09 ST0212 Homo sapiens cDNA, mRNA sequence.
LOCUS   BF375244   698 bp   mRNA   linear   EST 24-NOV-2000
RESULT 10

RESULT 11

```
Db   475 CGCAAACCCAGC-TCACTTGTACTGGAAAGGAGGAGGTGGAGCAACAAGGCTTATTGCA 533
         |||||||||||| |||||||||||||||||||||||||||||||||||||||||||| ||||
Qy   167 sLeuGlyGluHisAspGlySerLeuAspLeuArgSerArgArgSerValGlnGluGlyAs 187
Db   534 TTAGCGGGACATGATGATGGCAGTGTTGACTTGCGATCTAGACGATCAGTTCAAGAAGTAA 593
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   187 nProArgAla 190
Db   594 TCCAAGGGCC 603
```

```
LOCUS       BG219307      822 bp    mRNA    linear    EST 21-APR-2001
DEFINITION  RST39061 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG219307
VERSION     BG219307.1   GI:13745328
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 822)
  AUTHORS   Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
            Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.
            , Lerner,L., Costanzo,D., McElligott,K., Boozer,S., Mays,R., Smith
            ,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
            ,J., Danzig,J. and Ducar,M.
  TITLE     Creation of genome-wide protein expression libraries using random
            activation of gene expression
  JOURNAL   Nat. Biotechnol. 19 (5), 440-445 (2001)
  MEDLINE   21227151
COMMENT     Contact: Scott J. Cain
            Athersys, Inc.
            3201 Carnegie Ave, Cleveland, OH 44115, USA
            Tel: 216 431 9900
            Fax: 216 361 9596
            Email: scain@athersys.com
            High quality sequence stop: 473.
FEATURES             Location/Qualifiers
     source          1..822
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="Athersys RAGE Library"
                     /note="See 'Creation of Genome-wide Protein Expression
                     Libraries using Random Activation of Gene Expression',
                     Nature Biotechnology, In press. Note that even though the
                     cell type indicated is HT1080, since a random activation
                     method was used, these sequence tags are not necessarily
                     expressed in HT1080 under normal circumstances."
BASE COUNT      241 a    169 c    179 g    230 t    3 others
ORIGIN
```

Alignment Scores:
Pred. No.:
Score: 1.63e-76
Percent Similarity: 775.00
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 77.35%

Length: 822
Matches: 148
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

```
US-09-895-298A-83 (1-190) x BG219307 (1-822)

Qy   43 LeuLysProSerAlaAspCysGlyProPheArgGlyLeuProLeuPheIleHisSerIle 62
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   79 TTGAAGCCTTCAGCTGACTGTGGCCCTTTTCGAGGTCTGCCTCTTCATTCACTCCATC 138
Qy   63 TyrSerTrpIleAspThrLeuSerThrArgProGlyThrTyrLeuTrpValValTrpIleTyr 82
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  139 TACAGCTGGATCGACACACCCTAAGTACACGGCCTCGGCTACCCGTGGGTTGTTTGGATCTAT 198
Qy   83 ArgAsnLeuIleGlySerValHisPhePheLeuThrLeuValLeuIleIle 102
```

RESULT 12

```
Db   199 CGGAAACCTCATTGGAAGTGTGCAACTTCTTTTTTCAACCCTCACCCTATGTCGTAATCATC 258
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   103 ThrTyrLeuTyrTrpGlnIleThrGluGlyArgLysIleMetIleArgLeuLeuHisGlu 122
Db   259 ACCTATCTTPACTGGCAGATCACAGAGGGAAGGGAAGAAGGAGATTATGATAAGGCTGCTCCA 318
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   123 GlnIleIleAsnGluGlyLysAspLysMetPheLeuLeuIleGluLysLeuLysLysLeuGln 142
Db   319 CAGATCATTAATGAGGGCAAAGATAAAAATGTTCCTGATACAGAAAATGATCAAGCTCGAG 378
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   143 AspMetGlyLysLysAlaAsnProSerSerLeuValLeuGluArgArgGlyLeuValGlnGln 162
Db   379 GATATGGAGAAGAAAGCAAAACCCCAGCTCACTTGTTCTGGGAAAGGAGAGAGGTGGAGCAA 438
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   163 GlnGlyPheLeuHisLeuGlyLysHisLeuAspLeuGlySerLeuAspLeuArgSerArgSer 182
Db   439 CAAGGCTTTTGCATTTGGGGGAACATGATGGCAGTCTTCGACTTGCGATCTAGAGATCA 498
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   183 ValGlnGluGlyAsnProArgAla 190
Db   499 GTTCAAGAAGGTAAATCCAAGGGCC 522
```

```
LOCUS       BG198917      871 bp    mRNA    linear    EST 21-APR-2001
DEFINITION  RST18188 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG198917
VERSION     BG198917.1   GI:13720604
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 871)
  AUTHORS   Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
            Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.
            , Lerner,L., Costanzo,D., McElligott,K., Boozer,S., Mays,R., Smith
            ,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
            ,J., Danzig,J. and Ducar,M.
  TITLE     Creation of genome-wide protein expression libraries using random
            activation of gene expression
  JOURNAL   Nat. Biotechnol. 19 (5), 440-445 (2001)
  MEDLINE   21227151
COMMENT     Contact: Scott J. Cain
            Athersys, Inc.
            3201 Carnegie Ave, Cleveland, OH 44115, USA
            Tel: 216 431 9900
            Fax: 216 361 9596
            Email: scain@athersys.com
            High quality sequence stop: 498.
FEATURES             Location/Qualifiers
     source          1..871
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="Athersys RAGE Library"
                     /note="See 'Creation of Genome-wide Protein Expression
                     Libraries using Random Activation of Gene Expression',
                     Nature Biotechnology, In press. Note that even though the
                     cell type indicated is HT1080, since a random activation
                     method was used, these sequence tags are not necessarily
                     expressed in HT1080 under normal circumstances."
BASE COUNT      252 a    186 c    185 g    246 t    2 others
ORIGIN
```

Alignment Scores:
Pred. No.:
Score: 1.08e-74
Percent Similarity: 759.00
Best Local Similarity: 98.65%
Query Match: 98.65%
DB: 75.75%

Length: 871
Matches: 146
Conservative: 0
Mismatches: 2
Indels: 0
Gaps: 0

```
RESULT 13
BG982234/c

Db  537  GTTCAAGAAGGTAAATCCAAGGGCC 560
              ||||||||||||||||||||||||||
Qy  183  ValGlnGlyGlyAsnProArgAla 190

Db  477  CAGATCAATTAATCAGGGCAAAGATAAAAATGTTCCTGATAGAAAAATTGAATCAAGGCTGCAG 536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  357  CAAGGCTTTTCCATTGGGGACAGCAGTCTTGACTTGCGATCTAGAAGATCA
Qy  163  GlnGlyPheLeuHisLeuGlyGluHisAspGlySerLeuArgSerArgArgSer 182

Db  417  GATATGGGAGAAGAAAAGCAAAACCCCAGCTCACTTGTTCTGGAAAGGAGAGGTGGAGCAA 476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  143  AspMetGlyGluGluLysGlnAsnProSerSerLeuValLeuGluArgArgGluValGluGln 162

Db  357  GATAAGGGAGAAGAAAAGCAAAACCCCAGCTCACTTGTTCTGGAAAGGAGAGGTGGAGCAA
```

(Sequence listing continues — multiple query/database alignment blocks)

```
US-09-895-298A-83 (1-190) x BG198917 (1-871)

Qy   43  LeuLysProSerAlaAspCysGlyProPheArgGlyLeuProLeuPheHisSerIle 62
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  117  TTGAAGCCTTCAACTGACGTGTGGCCCTTTCGAGGTCTGCCTCTTCCATTCACTCCATC 176

Qy   63  TyrSerTrpIleAspThrLeuSerThrArgProGlyTyrLeuTrpValValTrpIleTyr 82
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  177  TACAGCTGGATCGACACCCTAAGTACACGACCCGGTTACCTGTGGGTTGTTTGGATCTAT 236

Qy   83  ArgAsnLeuIleGlySerValHisPhePheIleLeuThrIleLeuValLeuIleIle 102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  237  TGGAACCTCATTGGAAGTGTGCACTTCTTTTCATCCTCACCCTCATTGTGCTAATCATC 296

Qy  103  ThrTyrLeuTyrTrpGlnIleThrGluGlyAlaLysIleMetIleAlaArgLeuLeuHisGlu 122
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  297  ACCTATCTTTACTGGCAGATCACAGAGGAAGCAAAGATAATGATAAGGCTGCTCCATGAG 356

Qy  123  GlnIleLeuLeuAsnGluGlyLysAspLysMetPheLeuGlyLysProLeuSerLeuGln 142
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  357  CAGATCAATTAATCAGGGCAAAGATAAAAATGTTCCTGATAGAAAAATTGAATCAAGGCTGCAG 416
```

source
```
1. 447
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /clone_lib="CN0089"

   1. 447
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /note="organ: colon_normal; Vector: puc18; Site_1: SmaI;
   Site_2: SmaI; A mini-library was made by cloning products
   derived from ORESTES PCR (U.S. Letters Patent Application
   No. 196,716 - Ludwig Institute for Cancer Research)
   profiles into the pUC 18 vector. Reverse transcription of
   tissue mRNA and cDNA amplification were performed under
   low stringency conditions."
```

```
LOCUS       BG982234    447 bp    mRNA    linear    EST 12-JUN-2001
DEFINITION  CM4-CN0089-130201-723-e04 CN0089 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG982234
VERSION     BG982234.1    GI:14384969
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 447)
AUTHORS     Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
TITLE       Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE     20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?ti-CM4&t2-CM4-CN0089-
            130201-723-e04&t3=2001-02-13&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 25
            High quality sequence stop: 447.
FEATURES             Location/Qualifiers
```

BASE COUNT   114 a   107 c   95 g   131 t
ORIGIN

```
RESULT 14
BG198969

Db  326  CTCCATTGTGCTAATCATCACCTATCTTTACTGGCAGATCACAGAGGGAGGAAGAGGATTATG 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   97  LeuIleValLeuIleIleThrTyrLeuTyrTrpGlnIleThrGluGlyAlaLysIleMet 116

Db  386  TGGGCTGTTTTGGATCTATCGGAACCTCATTGGAAGTGTGCACTTCTTTTCATCCTCACC 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   77  TrpValValTrpIleTyrArgAsnLeuIleGlySerValHisPhePheIleLeuThr 96

Db  446  CTCTTCATTCACTCCATCTACAGCTGGATCGACACCCTAAGTACACGACCCGGCTACCTG 387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  117  LeuPheIleHisSerIleTyrSerTrpIleAspThrLeuSerThrArgProGlyTyrLeu 136

Db  266  ATAAGGCTGCTCCATGAGCAGATCAATTAATGAGGGCAAAGATAATGATAAGGCTGTTTTGGG 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  137  LysLeuLeuHisGluGlnIleLeuLeuAsnGluGlyLysAspLysMetPheLeuGlu 156

Db  206  AAATTGAATCAAGGCTGCAGGATATGGGAGAAGAAAAGCAAAACCCCAGCTGTCTGGAA 147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  157  ArgArgGluValGluGlnGlnGlnGlyPheLeuHisLeuGluGlyGluHisAspGlySerLeuAsp 176

Db  146  AGGAGAGAGGTGGAGCAACAAGGGCTTTTCCATTGGGGACAGCAGTCTTGAC 87
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  177  LeuArgSerArgArgSerValGlnGlyGlyAsnProArgAla 190

Db   86  TTGCGATCTAGAAGATCAGTTCAAGAAGGTAAATCCAAGGGCC 45
```

```
US-09-895-298A-83 (1-190) x BG982234 (1-447)
```

Alignment Scores:
```
Pred. No.:                      2.12e-67
Score:                          691.00
Percent Similarity:             99.25%
Best Local Similarity:          99.25%
Query Match:                    68.96%
DB:                             13
```

Length: 447
Matches: 133
Conservative: 0
Mismatches: 1
Indels: 1
Gaps: 0

```
LOCUS       BG198969    758 bp    mRNA    linear    EST 21-APR-2001
DEFINITION  RST18241 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG198969
VERSION     BG198969.1    GI:13720556
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 758)
AUTHORS     Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
            Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.
            , Lerner,L., Costanzo,D., McElligott,K., Boozer,S., Mays,R., Smith
            , E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
            ,J., Danzig,J. and Ducar,M.
TITLE       Creation of genome-wide protein expression libraries using random
            activation of gene expression
JOURNAL     Nat. Biotechnol. 19 (5), 440-445 (2001)
```

MEDLINE 21227151
COMMENT
  Contact: Scott J. Cain
  Athersys, Inc.
  3201 Carnegie Ave, Cleveland, OH 44115, USA
  Tel: 216 431 9900
  Fax: 216 361 9596
  Email: scain@athersys.com
  High quality sequence stop: 552.

FEATURES
source
  1. .758
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone_lib="taxon:9606"
  /clone_lib="Athersys RAGE Library"
  /cell_line="HT1080"
  /note="See 'Creation of Genome-wide Protein Expression
  Libraries using Random Activation of Gene Expression',
  Nature Biotechnology, in press. Note that even though the
  cell type indicated is HT1080, since a random activation
  method was used, these sequence tags are not necessarily
  expressed in HT1080 under normal circumstances."

BASE COUNT 172 a 193 c 163 g 229 t 1 others
ORIGIN

US-09-895-298A-83 (1-190) x BG198969 (1-758)

Alignment Scores:
Pred. No.:
Score: 2.32e-66
                            684.00
Percent Similarity: 94.90%   Length: 758
Best Local Similarity: 91.08%   Matches: 143
Query Match: 68.26%   Conservative: 6
DB: 12   Mismatches: 8
                                            Indels: 3
                                            Gaps: 0

```
Qy   1 MetMetAsnPheGlnProProSerLysAlaTrpAlaSerGlnMetMetThrPhePhe  20
         |||||||||||||||||||||||||||||||||||||||||||||
Db 284 ATGATGAAATTTCCAGCCTCCGAGCAAAGCCTGGCGGCCTCACAGATGATGACTTTCTTC 343

Qy  21 IlePheLeuLeuPhePheProSerPheThrArgGlyValLeuCysThrLeuAlaIleThrIle  40
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 344 ATCTTCTTGCTCTTTTTCCCATCCTTCACCGGGGTCTTGTGCACCCTGGCCATCACCATC 403

Qy  41 TrpArgLeuLysProSerAlaAspCysGlyProPheArgGlyLeuProLeuPheIleHis  60
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 404 TGGAGATTGAAGCCTTCAGCTGACTGTGGCCCTTTTCGAGGTGCTGCCTTTCTCATTCAC 463

Qy  61 SerIleTyrSerTrpIleAspThrLeuSerThrArgProGlyTyrLeuTrpTpValValTrp  80
         |||||||||||||||||||||||||||||||||||||||||||||||
Db 464 TCCATCTACAGCTGGATCGACACCCTRAGTACACGGCTCCTACCTGGGTTGGTTTGG 523

Qy  81 IleTyrArgAsnLeuIleGlyServalHisPhePheIleLeuThrLeuIleValLeu 100
         |||||||||||||||||||||||||||||||||||||||||||
Db 524 ATCTATCGGAACCTCATTGGAAGTGTGCACTTCTTTTCATTGTGCTA 583

Qy 101 IleLeuThrArgTyrLeuTrpTrpGlnIleLeuThrGlyLeuArgLysIleMetIleArgLeuLeu 120
         ||||||||||||||||||||||||||||||||||||||||
Db 584 ATCACACTTATCTTTACTGGCGAGATCACAGAGGGAAGGAAGATTATG 643

Qy 121 HisGluGlnIleIleAsnGluGlyLysAspLysLysSerMetPheLeuIleGluLys 140
         ::||||||||||||||||||::||||||||||||||||||
Db 644 CATGAGCAGATCATTAATGAGGGCAAAGATAAAAATGTTCCTGATAGAA 701

Qy 141 LeuGlnAspMetGlyLysLysLysAlaAsnProSerSerLeuValLeuGluArg 157
         |||||||||::|||:::|||||||||||||||::||:||:|
Db 702 CTGGAGGATATGGACAAGAAGAAAGCACACACCCCAGCTCA-CTAGTTCTGGAAAGA 751
```

SOURCE human.
ORGANISM Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 594)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
  Emmert-Buck, M.D., Ph.D.
  cDNA Library Preparation: Life Technologies, Inc.
  cDNA Library Arrayed by: Greg Lennon, Ph.D.
  DNA Sequencing by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/bbrp/image/image.html
  Seq primer: -40UP from Gibco
  /note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI;
  Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT;
  Average insert size 1.69 kb. Life Technologies catalog #:
  11549-011"

FEATURES
source
  1. .594
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone="IMAGE:2898861"
  /clone_lib="NCI_CGAP_Gas4"
  /tissue_type="poorly differentiated adenocarcinoma with
  signet ring cell features"
  /lab_host="DH10B"
  High quality sequence stop: 404.

BASE COUNT 161 a 139 c 108 g 186 t
ORIGIN

US-09-895-298A-83 (1-190) x AW469181 (1-594)

Alignment Scores:
Pred. No.:
Score: 4.78e-54
                            573.00
Percent Similarity: 99.12%   Length: 594
Best Local Similarity: 99.12%   Matches: 113
Query Match: 57.19%   Conservative: 0
DB: 10   Mismatches: 1
                                            Indels: 0
                                            Gaps: 0

```
Qy  77 TrpValValTrpIleTyrArgAsnLeuIleGlyServalHisPhePheIleLeuThr  96
         |||||||||||||||||||||||||||||||||||||||||||
Db 594 TGGGTGGTTGGAATCTATCGGAACCTCATTGGAAGTGTGCACTTCTTTTCATCCTCACC 535

Qy  97 LeuIleValLeuIleLeuThrArgTyrLeuTrpTrpGlnIleLeuThrGlyLeuArgLysIleMet 116
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 534 CTCATTGTGCTAATCACATCACCTTATCTTTACTGGCGAGATCACAGAGGGAAGGAAGATTATG 475

Qy 117 IleArgLeuLeuHisGluGlnIleIleAsnGluGlyLysAspLysLysSerMetPheLeuIleGlu 136
         |||||||||||||||||||||||||||||||||||||||||||||||
Db 474 ATAAGGCTGCTCCATGAGCAGATCATTAATGAGGGCAAAGATAAAAATGTTCCTGATAGAA 415

Qy 137 LysLeuIleLysLeuGlnAspMetGlyLysLysLysAlaAsnProSerSerLeuValLeuGlu 156
         ||||||||||||||||||||||||||||||||||||||||||||
Db 414 AAATTGATCAAGCTGCAGGATATGGAGAAGAAAAAGCAAACCCAGCTCACTGTTCTGGAA 355

Qy 157 ArgArgGluValGlnGlnGlyLysAsnProArgAla 176
         ||||||||||||||||||||||||||||||||||||
Db 354 AGGAGAGAGGTGGAGCAACAAGGCTTTTGCATTGGGGGAACATGATGATGGCAGTCTTGAC 295

Qy 177 LeuArgSerArgArgServalGlnGlnGlyLysAsnProArgAla 190
         |||||||||||||||||||||||||||||||||||||||||
Db 294 TTGCGATCTAGAAGATCAGTTCAAGAAGGTAATCCAAGGCC 253
```

Search completed: November 9, 2002, 06:12:16